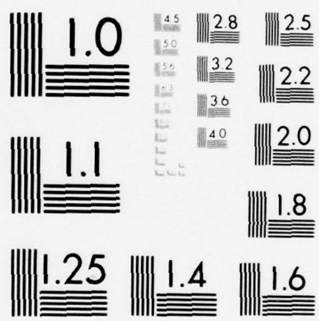
AD-A062 094 MARYLAND UNIV BALTIMORE COUNTY BALTIMORE DIV OF MATH--ETC F/G 12/1
OPTIMAL FILTERS FOR BILINEAR SYSTEMS WITH NILPOTENT LIE ALGEBRA--ETC(U)
OCT 78 S D CHIKTE, J TING-HO LO AFOSR-74-2671

UNCLASSIFIED MRR-78-13 AFOSR-TR-78-1505 NL

1 OF 1
AD
A062094

END
DATE
FILMED
3-79
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 78-1505

Mathematics Programs at U.M.B.C.

Mathematics Research Reports

LEVEL II

Division of Mathematics and Physics

UNIVERSITY OF MARYLAND
BALTIMORE COUNTY

Baltimore, Maryland 21228

78 12 04.069

(14) MRR-78-13

(6) Optimal Filters for Bilinear Systems
with Nilpotent Lie Algebras.

(12) 33 p.

by

(10) Shirish D. Chikte ~~and~~ James Ting-Ho Lo

Mathematics Research Report No. 78-13

(9)

(11) Oct[redacted]1978

(18) AFOSR (16) 2304 (17) A1

(19) TR-78-1505 (15) AFOSR-74-2671

Abstract

We consider a bilinear signal process driven by a Gauss-Markov process which is observed in additive, white, gaussian noise. An exact stochastic differential equation for the least squares filter is derived when the Lie algebra associated with the signal process is nilpotent. It is shown that the filter is also bilinear and moreover that it satisfies an analogous nilpotency condition. Finally, some special cases and an example are discussed, indicating ways of reducing the filter dimensionality.

Mailing Addresses

Shirish D. Chikte**
Department of Mathematics
University of Maryland Baltimore County
Baltimore, Maryland 21228
Telephone: (301) 455-2431

James Ting-Ho Lo
Department of Mathematics
University of Maryland Baltimore County
Baltimore, Maryland 21228
Telephone: (301) 455-2432

*This research was supported by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-74-2671C.

**Preferred address for correspondence.

LP

409 579

OPTIMAL FILTERS FOR BILINEAR SYSTEMS

WITH NILPOTENT LIE ALGEBRAS

1. INTRODUCTION

In recent years, detection, estimation and control of signal processes represented by bilinear systems has received some attention in the literature; see, e.g. articles [1], [2] and [3] and survey [4]. The principal motivation for studying this class of problems lies in its potential applications to a variety of practical areas such as inertial navigation, satellite attitude control and angle modulation.

We focus on least squares estimators in additive, white, gaussian noise environment. In [1], such estimators have been obtained in recursive and closed form under the assumption that Lie algebras associated with the signal process are abelian. In [3], the existence of such finite dimensional, recursive estimators has been established under the weaker requirement that these Lie algebras need only be nilpotent; no attempt, however, is made towards displaying the estimator equations themselves.

In the present paper, we derive explicitly, the finite dimensional, closed form, recursive filtering equations when the signal process satisfies a nilpotency condition, thus supplying a complete and constructive solution to this class of problems. In this process, we prove that the filter is bilinear as well and, moreover, that it also possesses analogous nipotent property. A number of interesting special cases are identified

in which the estimator can be alternatively realized via a linear filter followed by a nonlinear postprocessor, a structure that may prove advantageous from the viewpoint of practical implementation.

In the next section, we formulate the problem and present some mathematical preliminaries. The third section contains the main theorem on the properties of the optimal filter, and the final section deals with computational considerations.

2. PROBLEM STATEMENT AND PRELIMINARIES

Consider the following standard linear Ito models for the signal and observation processes respectively.

The Signal Model:

$$d\xi(t) = F(t)\xi(t)dt + Q^{1/2}(t)dw(t), \qquad t \geq 0 \tag{2.1}$$

The Observation Model:

$$dz(t) = H(t)\xi(t)dt + R^{1/2}(t)dv(t), \qquad t \geq 0 \tag{2.2}$$

where, $w(\cdot)$ and $v(\cdot)$ are standard N and P dimensional independent Wiener processes respectively, $\xi(t) \in \mathbb{R}^N$, $z(t) \in \mathbb{R}^P$, $\xi(0)$ is a zero mean, gaussian random vector independent of $w(\cdot)$ and $v(\cdot)$ processes and $F(\cdot)$, $Q^{1/2}(\cdot)$, $H(\cdot)$, $R^{1/2}(\cdot)$ are time-dependent matrices of appropriate dimensions, with $Q(t)$, $R(t)$ positive definite and continuously differentiable for all t.

Our interest in this paper centers on the least square estimation of nonanticipative, square integrable, nonrandom functionals on the Gauss-Markov process $\xi(\cdot)$ defined in (2.1), based on the observation process $z(\cdot)$ defined in (2.2). It is only natural to represent a functional of this type by a dynamical system driven by $\xi(\cdot)$.

Clearly, this problem fits the framework of the nonlinear filtering problem discussed by Kushner [5]. His solution, however, requires, in general, solving an infinite set of coupled stochastic differential equations. As a result, a number of approximation schemes have since been proposed to "close" this infinite set, notable among them: Extended Kalman filter, Symmetric Density Filter, Bass-Schwartz Filter, a cumulant discard hypothesis, fourth moment assumption and various other moment approximations; the interested render is referred reader is referred to Chapter 9 of [6] and reference therein. Each of these attempts have experienced varying degrees of success, depending on the specific practical application at hand. Besides, they have often lacked rigorous mathematical justification.

Our approach, here, is different, in that we seek to "close" the Kushner equations exactly by suitably restricting the class of nonlinear functionals to be estimated, thus obtaining an exact solution to the restricted problem. Moreover, we should like this restricted subset of functionals to be also "dense" in the class of L_2 - functionals, so that we would have essentially "solved" the global nonlinear problem as well. Analogous approach has in the past, been taken by Balakrishnan [7] and

more recently by Huang [8], Chapter V, although in a "static" framework, that is, nonrecursive estimation of a single random variable from fixed length data without the dynamical signal and noise framework.

In view of the above considerations and some results available (see [9] and [10]) on approximation of nonlinear systems with deterministic inputs, the class of bilinear systems seems to be the most promising subset on which one would like to focus attention. We, therefore, assume that the signal process $\{x(t)\}_{t \geq 0}$, $x(t) \in \mathbb{R}^M$, to be estimated, evolves according to the following bilinear dynamical equation:

$$dx(t) = Ax(t)dt + \sum_{i=1}^{N} B_i \xi_i(t)x(t)dt \tag{2.3}$$

where $A, B_1, \ldots, B_N$ are $M \times M$ constant matrices, and $x(0)$ is independent of $\xi(0)$ and $w(\cdot)$ and $v(\cdot)$ processes. With this model, we seek a finite dimensional stochastic differential equation for computing $x(t/t) \triangleq E[x(t)/z^t]$.

Besides the aforementioned mathematical considerations, the above signal model has strong justification on physical grounds as well. As discussed e.g. in [1] and [11], (the state transition matrices of) bilinear systems, evolving on Lie groups can perfectly represent certain types of motion such as rotation of rigid bodies.

It may be appropriate at this point to recall some pertinent definitions and facts from the theory of Lie algebras and Lie groups. Further details may be found, for example, in [12]. Let L denote a Lie algebra and

$G_L \triangleq e^L$, the associated Lie group. It is easy to verify that the set

$$L_1 \triangleq [L,L] \triangleq \{[A,B] \mid A,B \in L\} \tag{2.4}$$

is an (ideal) subalgebra, where $[A,B] \triangleq AB-BA$, is the Lie bracket operation. Now define analogously the following two series of decreasing, nested subalgebras, recursively as,

$$L^k \triangleq [L,L^{k-1}] \triangleq \{[A,B] \mid A \in L, B \in L^{k-1}\} \qquad k = 2,3,\ldots, \text{ with } L^1 = L_1 \tag{2.5}$$

and

$$L_k \triangleq [L_{k-1}, L_{k-1}] \triangleq \{[A,B] \mid A,B \in L_{k-1}\}, \quad k = 2,3,\ldots \tag{2.6}$$

<u>Definition 2.1</u>: The Lie algebra L (and the Lie group G_L) is said to be

a) <u>abelian</u> if $L_1 = \{0\}$

b) <u>nilpotent</u> if there is an integer K such that $L^K = \{0\}$

c) <u>solvable</u> if there is an integer K such that $L_K = \{0\}$.

Analogous definitions can be made with respect to an <u>associative</u> algebra as well, merely by replacing Lie bracket operation by ordinary matrix multiplication in the above discussion. It then follows directly from the definitions that a) $\Rightarrow$ b) $\Rightarrow$ c) and that for a Lie algebra to possess one or more of these properties it is sufficient that the smallest matrix algebra containing the Lie algebra also possess the corresponding property.

Coming back to our estimation problem summarized in the signal and noise models (2.1), (2.2) and (2.3), it was shown in [1], that if the Lie algebra generated by the matrices $A, B_1, \ldots, B_N$, denoted by $\{A, B_1, \ldots, B_N\}_L$ is abelian, the estimator for $x(\cdot)$ consists of a linear filter and a nonlinear postprocessor. In [3], the abelian type condition was replaced by a weaker nilpotency condition. However, [3] establishes merely the existence of a finite dimensional recursive filter. In the next section, we derive explicitly under the above Lie-algebraic nilpotency condition — the stochastic differential equations of the filter and futhermore explore its algebraic properties as well. A perliminary analysis of this type for the much simpler special case of associate algebraic nilpotency may be found in [13].

3. PROPERTIES OF THE NON-LINEAR FILTER

We begin with some lemmas which will be heavily used in the proof of the theorem.

Lemma 1 (Canonical Nilpotent Form):

Every nilpotent matrix Lie algebra can be converted, by a similarity transformation, into its canonical Lie algebra consisting of block diagonal matrices wherein each diagonal block is triangular with equal elements on the diagonal.

Proof: See Sagle and Walde [12] pp. 224-227.

Lemma 2 (Exponential Formula):

Let $\mathcal{L}$ be a Lie algebra of matrices with $\left\{H_i\right\}_{i=1}^{N}$ as its basis. Then for any pair $A,B \in \mathcal{L}$ we have

$$e^{At}Be^{-At} = \sum_{i=1}^{N} g_i(t)H_i \tag{3.1}$$

where $\left\{g_i(\cdot)\right\}_{i=1}^{N}$ are analytic functions.

Proof: See lemmas (i) and (ii) of Wei and Norman [14].

Lemma 3:

Consider the signal and observation models of (2.1) and (2.2) respectively. Define a (vector) process $\{y(t)\}_{t \geq 0}$ by

$$\left.\begin{array}{l} dy(t) = Dy(t)dt + \sum_{i=1}^{N} E_i\xi_i(t)y(t) \text{ where: } D, \left\{E_i\right\}_{i=1}^{N} \text{ are} \\ \text{matrices of appropriate dimension, } y(0) \text{ is} \\ \text{independent of } \xi(0), w(\cdot) \text{ and } v(\cdot) \text{ processes.} \end{array}\right\} \tag{3.2}$$

Then $\hat{y}(t/t) \triangleq E[y(t)/z^t]$ satisfies the following stochastic differential equation:

$$\left.\begin{array}{l} d\hat{y}(t/t) = D\hat{y}(t/t)dt + \sum_{i=1}^{N} E_i \cdot E^t\left[\xi_i(t)y(t)\right]dt + \left\{E^t\left[y(t)\xi^T(t)\right]\right. \\ \qquad \left. - \hat{y}(t/t)\hat{\xi}^T(t/t)\right\} H^T(t)R^{-1}(t)d\nu(t) \\ \hat{y}(0/0) = E[y(0)] \end{array}\right\} \tag{3.3}$$

where $E^t[\cdot] \triangleq E[\cdot|z^t] \triangleq E[\cdot|\{z(\tau)|\ 0 \leq \tau \leq t\}]$

$d\nu(t) = dz(t) - H(t)\hat{\xi}(t/t)dt.$

Proof: Apply the Kushner nonlinear filtering equations [5] to the signal process $(y^T(\cdot), \xi^T(\cdot))^T$ with $z(\cdot)$ as the observation process.

Lemma 4:

Let $x = (x_o, x_1, \ldots, x_n)^T$ be a gussian random vector with mean vector $m = (m_o, m_1, \ldots, m_n)^T$ and covariance matrix $P = [P_{ij}]_{i,j=0}^{n}$. We then have the following relation:

$$E\left[e^{x_o}\prod_{i=1}^{n} x_i\right] = \begin{cases} (m_1 + P_{o1}) \cdot E\left[e^{x_o}\prod_{i=2}^{n} x_i\right] + E\left[e^{x_o}\sum_{j=2}^{n} P_{1j} E\left(\prod_{\substack{i=2\\ i\neq j}}^{n} x_i\right)\right], & n > 1 \\ (m_1 + p_{o1})\, E\left[e^{x_o}\right], & n = 1 . \end{cases} \tag{3.4}$$

Proof: We shall indicate only the main steps of the proof leaving the purely algebraic manipulations to the reader.

$$E\left[e^{x_o}\prod_{i=1}^{n} x_i\right] = \int_{\mathbb{R}^{n+1}} (2\pi|P|)^{-(n+1)/2} \exp - \tfrac{1}{2}(x-m)^T P^{-1}(x-m) \cdot e^{x_o}\prod_{i=1}^{n} x_i \, dx$$

(where $|\cdot| \triangleq \det(\cdot)$)

$$= (2\pi|P|)^{-(n+1)/2} \cdot e^{m_o + \frac{1}{2Q_{oo}}} \int_{\mathbb{R}^{n+1}} \prod_{i=1}^{n} x_i \cdot$$

$$\exp - \frac{1}{2}\left[(x_o - m_o - \frac{1}{Q_{oo}})^2 \cdot Q_{oo} + \sum_{\substack{i,j=0\\ i=j\neq 0}}^{n} Q_{ij}(x_i - m_i)(x_j - m_j)\right] dx$$

(upon completing the square for the term $i = j = 0$ and using the notation $Q_{ij} \triangleq i,j^{th}$ element of $P^{-1} \triangleq Q$),

$$= e^{(m_o + \frac{1}{2}P_{oo})} \cdot (2\pi|P|)^{\frac{-(n+1)}{2}} \int_{\mathbb{R}^{n+1}} dx \prod_{i-1}^{n} x_i \exp - \frac{1}{2}\Big[(x_o - m_o - \frac{1}{Q_{oo}})^2 Q_{oo}$$

$$+ \sum_{\substack{i,j=0 \\ i=0, j\neq 0}}^{n} Q_{oj}(x_o - m_o - \frac{1}{Q_{oo}})(x_j - m_j) + \sum_{\substack{i,j=0 \\ i\neq 0, j=0}}^{n} Q_{io}(x_i - m_i)(x_o - m_o - \frac{1}{Q_{oo}})$$

$$+ \sum_{i,j=1}^{n} Q_{ij}(x_i - m_i)(x_j - m_j)\Big] \cdot \exp\Big[-\sum_{i=1}^{n}\Big(\frac{Q_{oi}}{Q_{oo}}(x_i - m_i) - \frac{P_{oi}}{P_{oo}} \cdot \frac{C_{oi}}{C_{oo}}\Big)\Big]$$

(where, $C_{ij} \triangleq$ cofactor of P_{ij})

$$= e^{(m_o + \frac{1}{2}P_{oo})} |2\pi P_*|^{-\frac{n}{2}} \int_{\mathbb{R}^n} \prod_{i=1}^{n} x_i \exp - \frac{1}{2}\Big[(x_* - m_*)^T P_*^{-1}(x_* - m_*)\Big]$$

$$\exp\Big[\sum_{i=1}^{n} \frac{C_{oi}}{|P_*|}(P_{oi} - x_i + m_i)\Big] dx_*$$

(integrating with respect to x_o, using the notation $P_* \triangleq [P_{ij}]_{i,j=1}$, $x_* = (x_1, \ldots, x_n)^T$ and the facts that $\frac{Q_{ij}}{Q_{i'j'}} = \frac{C_{ij}}{C_{i'j'}}$ for choice of integers i, j, i', j' and $C_{oo} = |P_*|$,

$$= e^{(m_o + \frac{1}{2}P_{oo})} \cdot |2\pi P_*|^{-\frac{n}{2}} \int_{\mathbb{R}^n} \prod_{i=1}^{n} x_i \exp - \frac{1}{2}\Big[(x_* - m_* - P_o^*)^T P^{-1}(x_* - m_* - P_o^*)\Big] dx$$

(combining the exponents, and noting the fact that $C_{o2} = \sum_{j=1}^{n} P_{oj} C_{ij}^*$, with with $C_{ij}^* \triangleq$ cofactor of P_{ij} <u>in P_*</u>. Also $P_o^* \triangleq [P_{o1}\ P_{o2} \ldots P_{on}]^T$.)

We have thus arrived at the following conclusion:

$$E\Big[e^{x_o} \prod_{i=1}^{n} x_i\Big] = e^{(m_o + \frac{1}{2}P_{oo})} \cdot E\Big[\prod_{i=1}^{n} Y_i\Big] \tag{3.5}$$

where $Y = (Y_1 \ldots Y_n)^T$ is a gaussian random vector with mean vector $(m_1 + P_{o1}, \ldots, m_n + P_{on})^T$ and P_* as the covariance matrix.

Now, a standard moment theorem for gaussian vectors (see e.g. [15]) yields

$$E\left[\prod_{i=1}^{n} Y_i\right] = (m_1 + P_{o1})E\left[\prod_{i=2}^{n} Y_i\right] + \sum_{j=2}^{n} P_{1j}E\left[\prod_{\substack{k=2 \\ k\neq j}}^{n} Y_k\right] . \tag{3.6}$$

Combining (3.5) and (3.6), and using repeatedly identity analogous to (3.5) for vector x with reduced dimensionality gives (3.4), completing the proof.

Q.E.D.

We are now in a position to state and prove the main theorem of this paper.

Theorem 5: Consider the signal process $\{x(t)\}_{t \geq 0}$ evolving according to (2.3) and (2.1) and the observation process $\{z(t)\}_{t \geq 0}$ as in (2.2). Suppose that the Lie algebra $\mathcal{L}$ generated by the set of matrices

$$\{Ad_A^k(B_i) \mid i = 1,2,\ldots, N; \quad K = 0,1,2,\ldots\}$$

is nilpotent, with dimension N^* and order of nilpotency N_*, where

$$Ad_A^o(B_i) \triangleq B_i$$

and, for $k = 1,2,\ldots$

$$Ad_A^k(B_i) \triangleq A \cdot Ad_A^{k-1}(B_i) - Ad_A^{k-1}(B_i) \cdot A .$$

Then the least squares filtered estimate $\hat{x}(t/t) \triangleq E\left[x(t)/\{z(\tau)\}_{0 \leq \tau \leq t}\right]$ can be obtained from the finite dimensional, bilinear, stochastic differential equation of the following form:

$$
\left.
\begin{aligned}
&d\hat{x}^*(t/t) = \left[A^*(t)dt + \sum_{i=1}^{N} B_i^*(t)\hat{\xi}_i(t/t) + \sum_{i=1}^{N} C_i^*(t)d\mu_i(t)\right]\hat{x}^*(t/t),\\
&\quad \hat{x}^*(\cdot/\cdot) \in \mathbb{R}^{M^*}, \; M^* \le M \cdot \left(\frac{(N^2N^*)^M - 1}{N^2N^* - 1}\right)\\
&\quad \hat{x}_i^*(0/0) = \begin{cases} E[x_i(0)], & i \le M \\ 0, & i > M \end{cases}\\
&\quad \hat{x}(t/t) = L(t)\hat{x}^*(t/t)\\
&\text{where}\\
&\quad \mu(t) \triangleq \int_0^t H^T(\tau)R^{-1}(\tau)d\nu(\tau), \quad \text{the modified innovations process,}
\end{aligned}
\right\} \quad (3.7)
$$

where, $\hat{\xi}(\cdot/\cdot)$ is obtained from the standard Kalman-Bucy filter (see [16]), $L(\cdot)$ is an $M \times M^*$ and $A^*(\cdot)$, $\left\{B_i^*(\cdot)\right\}_{i=1}^{N^*}$, $\left\{C_i^*(\cdot)\right\}_{i=1}^{N^*}$ are $M^* \times M^*$ matrix valued deterministic time functions such that the Lie algebra $\mathcal{L}$ generated by the set of matrices

$$
\left\{ \mathrm{Ad}^k_{\left[\mathrm{Ad}^\ell_{C_i^*(t)}(A^*(t))\right]}(B_j^*(t)) \;\middle|\; i = 1,2,\ldots,N;\; j = 1,2,\ldots,N;\; k,\ell = 0,1,\ldots;\; t \ge 0 \right\}
$$

is nilpotent with N_* as the order of nilpotency.

<u>Proof</u>: Let $\left\{H_i\right\}_{i=1}^{N^*}$ be a basis of $\mathcal{L}$ and $\mathcal{L}_c$ the canonical form of $\mathcal{L}$. Hence, by Lemma 1, there exists a (nonsingular) matrix S such that the set $\left\{H_i^*\right\}_{i=1}^{N^*}$ with $H_i^* \triangleq SH_iS^{-1}$, $i = 1,\ldots, N^*$ is a basis of $\mathcal{L}_c$ and $H_i^* = \mathrm{diag}\left\{{}^1H_i^*, {}^2H_i^*, \ldots, {}^\ell H_i^*\right\}$ for all $i = 1,\ldots, N^*$, where, each diagonal block ${}^kH_i^*$, $i = 1,\ldots, N^*$, $k = 1,\ldots,\ell$ is $M \times M$, $\sum_{k=1}^{\ell} M_k = M$,

and of the following form:

$$ {}^{k}H_i^* = h_{ii}^{k} I_{M_k} + {}^{k}\bar{H}_i^* \;,\; h_{ii}^{k} \in \mathbb{R},\; {}^{k}\bar{H}_i^* \quad \underline{\text{strictly upper triangular}}. \tag{3.8} $$

Now, consider the transformation

$$ y(t) = \left[Se^{-At}\right]x(t) \quad . \tag{3.9} $$

Using the above observation and Lemma 2, the transformed version of (2.3) can be written as,

$$ dy(t) = \left[\sum_{i=1}^{N^*} H_i^* \xi_i^*(t)\right] y(t)dt \tag{3.10} $$

where, $\xi^*(t) = D(t)\xi(t)$, *and* $D(\cdot)$ *is a deterministic* $N^* \times N$ *matrix* valued (analytic) time function. But, note that the dynamical system (3.10) is in a "decoupled" form and hence is the direct sum of the ℓ "subsystems"

$$ dy^k(t) = \left[\sum_{i=1}^{N^*} {}^{k}H_i^* \xi_i^*(t)\right] y^k(t), \tag{3.11} $$

where, $y^k(\cdot)$ is the M_k vector as follows

$y^k = \left(y_{M_{k-1}+1}, y_{M_{k-1}+2}, \ldots\ldots y_{M_{k-1}+M_k}\right)^T$, (with $M_o \triangleq 0$). We have thus effectively segregated the filtering problem into ℓ independent subproblems. Hence (except for a possible increase in filter dimensionality) we may assume without loss of generality that $\underline{\ell = 1}$, so that the original system has no nontrivial subsystems beside itself.

Now, apply the filtering equation (3.3) of Lemma 3 to the system (3.10) with

$$H_k^* = \begin{pmatrix} h_k & h_{12}^k & \cdots & h_{1M}^k \\ & h_k & \ddots & \vdots \\ & & \ddots & h_{M-1,M}^k \\ 0 & & & h_k \end{pmatrix}, \; k=1,2,\ldots,N^* . \tag{3.12}$$

First note that the solution to this system (for $\sigma \geq 0$) can be explicitly written as

$$y_q(\xi_o) = \begin{cases} e^{\xi_o^*(\sigma_o)} \cdot y_q(0) & q = M \\ e^{\xi_o^*(\sigma_o)} \cdot \left\{ \sum_{u=q+1}^{M} \sum_{s=1}^{u-q} \sum_{m_1,\ldots,m_s=1}^{N^*} \sum_{q=i_1<j_1=i_2<\cdots j_s=u} y_u(0) \int_0^{\sigma_o}\int_0^{\sigma_1}\cdots\int_0^{\sigma_{M-q-1}} \prod_{\ell=1}^{s} \left(h_{i_\ell,j_\ell}^{m_\ell} \xi_{m_\ell}^*(\sigma_\ell) d\sigma_\ell \right) \right\}, & q < M \end{cases} \tag{3.13}$$

where

$$\xi_o^*(\sigma_o) = \sum_{k=1}^{N^*} h_k \int_0^{\sigma_o} \xi_k^*(\tau) d\tau. \tag{3.14}$$

Now, using identify (3.4)—with conditional expectations—of Lemma 4, application of (3.3) to (3.10) yields:

$$d\hat{y}(t/t) = \left[\sum_{j=1}^{N} H_j^B(t)\hat{\xi}_j(t/t) + P_o(t)\right]\hat{y}(t/t)dt + \sum_{k=1}^{N}\sum_{i=1}^{N^*}\sum_{j=1}^{N} H^A_{k,i,j}(t)\hat{y}^{(k,i,j)}(t/t)dt$$

$$+ \left\{\left[\sum_{i=1}^{N^*}\sum_{j=1}^{N} D_{ij}(t)\hat{y}^{(1,i,j)}(t/t) + P_o(t)\hat{y}(t/t)\right], \left[\sum_{i=1}^{N^*}\sum_{j=1}^{N} D_{ij}(t)\hat{y}^{(2,i,j)}(t/t)\right.\right.$$

$$\left.\left. + P_o(t)\hat{y}(t/t)\right], \cdots \left[\sum_{i=1}^{N^*}\sum_{j=1}^{N} D_{ij}(t)\hat{y}^{(N,i,j)}(t/t) + P_o(t)\hat{y}(t/t)\right]\right\}$$

$$\cdot\ H^T(t)R^{-1}(t)d\nu(t)$$

$$\hat{y}(0/0) = E[y(0)], \tag{3.15}$$

where, the deterministic matrix valued time functions $\left\{H_j^B(\cdot)\right\}_{j=1}^{N}$, $P_o(\cdot)$ and $\left\{H^A_{k,i,j}(\cdot)\ \middle|\ k = 1,\ldots, N^*;\ i,j = 1,\ldots,N\right\}$ belong to the linear manifold $\mathcal{M}\left\{H_j^*\right\}_{j=1}^{N^*}$ generated by the set $\left\{H_j^*\right\}_{j=1}^{N^*}$ and can be computed from knowledge of the covariance matrix $P(\cdot)$ of the $\xi(\cdot)$ process and the matrix $D(\cdot) \triangleq [D_{ij}(\cdot)]$. Further, the "supplementary state vectors"

$\left\{y^{(k,i,j)}(\cdot)\ \middle|\ j,k = 1,\ldots,N;\ i = 1,\ldots,N^*\right\}$ appearing in (3.15) are defined as follows:

$$y^{(k,i,j)}(t) = \left[y_1^{(k,i,j)}(t), y_2^{(k,i,j)}(t), \ldots\ldots, y_M^{(k,i,j)}(t)\right]^T,$$

with

$$y_q^{(k,i,j)}(\sigma_o) = \begin{cases} 0, & q = M \\ \displaystyle\sum_{u=q+1}^{M}\sum_{s=1}^{u-q}\sum_{m_1,\ldots,m_s=1}^{N^*}\sum_{q=i_1<j_1=i_2\cdots<j_s=u} & \end{cases} \tag{3.16}$$

$$\sum_{\substack{r=1\\ m_r=k}}^{M-q} e^{\xi_o^*(\sigma_o)} y_u(0)\int_0^{\sigma_o}\int_0^{\sigma_1}\cdots\int_0^{\sigma_{M-q-1}} P_{ij}(\sigma_o,\sigma_r)h^{m_r}_{i_r,j_r}\,d\sigma_r\prod_{\substack{\ell=1\\ \ell\neq r}}^{s}\left\{h^{m_\ell}_{i_\ell,j_\ell}\xi^*_{m_\ell}(\sigma_\ell)d\sigma_\ell\right\},$$

$q < M$

and we have used the well-known fact that the conditional covariance matrix given by

$$P(\sigma, t) \triangleq E\left[(\xi(\sigma) - \hat{\xi}(\sigma/t)(\xi/\sigma) - \hat{\xi}(\sigma/t)^T/z^t\right] \tag{3.17}$$

is nonrandom for $\sigma \leq t$.

A direct differentiation now reveals that the vector of augmenting states

$$^1y(\cdot) \triangleq \left[\overset{(1,1,1)^T}{y(\cdot)} \quad \overset{(1,1,2)^T}{y(\cdot)} \quad \ldots\ldots \quad \overset{(N,N^*,N)^T}{y(\cdot)}\right]^T \in \mathbb{R}^{(MN^2N^*)} \tag{3.18}$$

satisfies a differential equation of the following form:

$$\left.\begin{aligned} d\,^1y(t) &= \left[\alpha^1(t) + \sum_{i=1}^{N} \gamma_i^1(t)\xi_i(t)\right]{}^1y(t)dt + \beta^1(t)y(t)dt \\ ^1y(0) &= 0 \end{aligned}\right\} \tag{3.19}$$

where we have used the formula (see e.g. [17]) frequently used in fixed point smoothing, viz;

$$\frac{\partial P(t,\sigma)}{\partial t} = \left[F(t) - P(t)H^T(t)R^{-1}(t)H(t)\right]P(t,\sigma) \tag{3.20}$$

and the $M \times M$ blocks of matrices $\alpha^1(\cdot)$, $\beta^1(\cdot)$ and $\left\{\gamma_i^1(\cdot)\right\}_{i=1}^{N}$ belong respectively to the

$$\mathcal{M}\left\{T_1^M I_M,\ T_1^M H_1^*,\ T_1^M H_2^*, \ldots,\ T_1^M H_{N^*}^*\right\},\ \mathcal{M}\left\{T_1^M H_1^*, T_1^M H_2^*, \ldots, T_1^M H_{N^*}^*\right\} \text{ and}$$

$$\mathcal{M}\left\{H_1 T_1^M,\ H_2 T_1^M, \ldots,\ H_{N^*} T_1^M\right\},$$

$$T_i^n \triangleq \left[\begin{array}{c|c} I_{n-i} & 0 \\ \hline 0 & 0 \end{array}\right] , \quad 1 \leq i \leq n \tag{3.21}$$

Furthermore, $\left\{\gamma_i^1(\cdot)\right\}_{i=1}^N$ are block diagonal matrices with identical diagonal blocks given by $H_i^B(\cdot)T_1^M$, $i = 1,\ldots, N$. For future reference, we also note that since $y_M^{(k,i,j)} \triangleq 0$, $H_{k,i,j}^A(\cdot) \in \mathcal{M}\left\{H_j^B(\cdot)T_1^M\right\}_{j=1}^{N^*}$.

We may now again apply lemmas 3 and 4 to (3.19) and obtain the following differential equation for ${}^1\hat{y}(t/t)$.

$$\begin{aligned} d^1\hat{y}(t/t) = {} & \left[\alpha^1(t) + \sum_{i=1}^{N} \gamma_i^B(t)\hat{\xi}_i(t/t) + P_o(t)\right]{}^1\hat{y}(t/t)dt \\ & + \beta^1(t)\hat{y}(t/t)dt + \sum_{k=1}^{N}\sum_{i=1}^{N^*}\sum_{j=1}^{N}\gamma_{k,i,j}^A(t)\,{}^1\hat{y}^{(k,i,j)}(t/t) \\ & + \left[P_o(t)\,{}^1\hat{y}(t/t) + \sum_{i=1}^{N^*}\sum_{j=1}^{N} D_{ij}(t)\,{}^1\hat{y}^{(1,i,j)}(t/t)\ ,\ P_o(t)\,{}^1\hat{y}(t/t)\right. \\ & + \sum_{i=1}^{N^*}\sum_{j=1}^{N} D_{ij}(t)\,{}^1\hat{y}^{(2,i,j)}(t/t)\ ,\ldots P_o(t)\,{}^1\hat{y}(t/t) \\ & \left. + \sum_{i=1}^{N^*}\sum_{j=1}^{N} D_{ij}(t)\,{}^1\hat{y}^{(N,i,j)}(t/t)\right]H^T(t)R^{-1}(t)d\nu(t) \end{aligned}$$

$${}^1\hat{y}(0/0) = 0 \ . \tag{3.22}$$

The "new" set of supplementary states appearing in (3.22) is in turn given by

$$ {}^{1}y^{(k,i,j)}(t) \triangleq \left[y^{(k,i,j)^T}_{(1,1,1)}(t),\ y^{(k,i,j)^T}_{(1,1,2)}(t), \ldots\ldots,\ y^{(k,i,j)^T}_{(N,N^*,N)} \right]^T, $$

$$ j,k = 1,\ldots, N;\ i = 1,\ldots, N^* $$

$$ y^{(k,i,j)}_{(k',i',j')}(t) \triangleq \left[y_1{}^{(k,i,j)}_{(k',i',j')}(t),\ y_2{}^{(k,i,j)}_{(k',i',j')}(t), \ldots\ldots,\ y_M{}^{(k,i,j)}_{(k',i',j')}(t) \right]^T $$

$$ y^{(k,i,j)}_{(k',i',j')}(\cdot) = \begin{cases} 0 & q \geq M-1 \\ \displaystyle\sum_{u=q+1}^{M} \sum_{s=1}^{u-q} \sum_{m_1,\ldots,m_s=1}^{N^*} \sum_{q=i_1 < j_1 = i_2 < \cdots j_s = u} \end{cases} $$

$$ \sum_{\substack{r_1, r_2 = 1 \\ r_1 \neq r_2 \\ m_{r_1} = k \\ m_{r_2} = k'}}^{M-q} e^{\xi_o^*(\sigma_o)} y_u(o) \int_0^{\sigma_o} \int_0^{\sigma_1} \cdots \int_0^{\sigma_{M-q-1}} P_{ij}(\sigma_o, \sigma_{r_1}) P_{i'j'}(\sigma_o, \sigma r_2) $$

$$ h_{i_{r_1}, j_{r_1}}^{m_{r_1}}\ h_{i_{r_2}, j_{r_2}}^{m_{r_2}}\ d\sigma_{r_1}, d\sigma_{r_2} \prod_{\substack{\ell=1 \\ \ell \neq r_1, r_2}}^{S} \left(h_{i_\ell, j_\ell}^{m_\ell}\ \xi_{m_\ell}^*(\sigma_\ell) d\sigma_\ell \right) \tag{3.23} $$

Now, the "new" augmenting state vector

$$ {}^{2}y(t) \triangleq \left[{}^{1}y^{(1,1,1)^T}(t),\ {}^{1}y^{(1,1,2)^T}(t), \ldots\ldots\ {}^{1}y^{(N,N^*,N)^T}(t) \right]^T \in \mathbb{R}^{MN^4 N^{*2}} \tag{3.24} $$

is easily seen to satisfy the following differential equation analogous to (3.19)

$$ d^2y(t) = \left[\alpha^2(t) + \sum_{i=1}^{N} \gamma_i^2(t) \xi_i(t) \right] {}^2y(t)dt + \beta^2(t)\, {}^1y(t)dt, \quad {}^2y(0) = 0 \tag{3.25} $$

where the $M \times M$ blocks of matrices $\alpha^2(\cdot)$, $\beta^2(\cdot)$ and $\left\{\gamma_i^2(\cdot)\right\}_{i=1}^{N}$ now belong to $\mathcal{M}\left\{T_2^M I_M,\ T_2^M H_1^*, \ldots,\ T_2^M H_{N^*}^*\right\}, \mathcal{M}\left\{T_2^M H_1^*,\ T_2^M H_2^*, \ldots,\ T_2^M H_{N^*}\right\}$ and $\mathcal{M}\left\{H_1 T_2^M, H_2 T_2^M, \ldots,\ H_{N^*} T_2^M\right\}$. Also $\gamma_i^2(\cdot)$, $i = 1, \ldots, N$ are diagonal with blocks given by $H_j^B(\cdot) T_2^M$, and $\gamma_{k,i,j}^A(\cdot) \in \mathcal{M}\left\{H_j^B(\cdot) T_2^M\right\}_{j=1}^{N^*}$.

It is now clear how the above process can be iterated. At exactly the M^{th} application of Kushner equations we find that $\gamma_i^{M-1} = 0$, $i = 1,2,\ldots, N$ so that no new supplementary states appear, resulting in closing the chain of coupled nonlinear filtering equations. Define

$x^*(t) \triangleq \left[y^T(t)\ {}^1y^T(t), \ldots,\ {}^{(M-1)}y^T(t)\right]^T$. We see that the dimension of $x^*(\cdot)$ is given by $M + M(N^2N^*) + M(N^2N^*)^2 + \ldots + M(N^2N^*)^{M-1} =$

$M\left(\frac{(N^2N^*)^{M-1}}{N^2N^*-1}\right)$. The filter of the form (3.7) results upon rewriting the innovations term in the standard bilinear format.

It only remains to prove the nilpotency of $\mathcal{L}^*$. Towards this end, we note the following features of (3.7)

(i) Each vector ${}^jy(t), 1 \le j \le M-1$ is "coupled" at most to its "adjacent" vectors ${}^{j-1}y(t)$ and ${}^{j+1}y(t)$ (through $A^*(\cdot)$ and $C_i^*(\cdot)$, $i = 1, \ldots, N$ matrices.)

(ii) $\frac{(N^2N^*)^{I-1}}{N^2N^*-1} < i \le \frac{(N^2N^*)^{I}-1}{N^2N^*-1}$, $I = 1, \ldots, M$.
Then the i^{th} row of $M \times M$ blocks of $A^*(\cdot)$, $\left\{B_j^*\right\}_{j=1}^{N}$ and $\left\{C_j^*\right\}_{j=1}^{N}$ are elements of
$\mathcal{M}\left\{T_I^M I_M,\ T_I^M H_1^*, \ldots,\ T_I^M H_{N^*}^*,\ H_1^*\ T_I^M, \ldots,\ H_{N^*}^* T_I^M\right\}$ with $(i,i)^{th}$ block of B_j^* equal to $H_j^B T_I^M$.

(iii) $\left\{B_i^*\right\}_{i=1}^{N^*}$ are <u>block diagonal</u> — hence, in nilpotent canonical form — while $\left\{C_i^*\right\}$ is <u>block triangular</u> with each $M \times M$ block being a multiple of $T_I^M I_M$, $1 \leq I \leq M$.

Keeping in mind the above observations and carrying out the Lie bracket operations blockwise, we find that the matrices $\bar{A}_{\ell,i}(t) \triangleq \mathrm{Ad}^{\ell}_{C_i^*(t)}(A^*(t))$, $i = 1,\ldots, N$; $\ell = 0,1,\ldots$ inherit all the properties of $A^*(t)$ noted above so that all the $M \times M$ block of matrices $\bar{B}_{k,j,\ell,i}(t) \triangleq \mathrm{Ad}^k_{\bar{A}_{\ell,i}}(t)\,(B_j^*(t))$, $k = 1,2,\ldots$; $j = 1,\ldots, N$ are strictly triangular. Since, as noted above, $\left\{B_i^*(t)\right\}_{i=1}^{N}$ are themselves in nipotent canonial form, the desired conclusion can be verified simply by carrying out blockwise the Lie bracket operations required in the definition (2.1b) of nilpotency.

<u>Q.E.D.</u>

We thus see that the optimal filter structure (3.7) is similar to that of the signal model (2.5) in that (i) it is bilinear in both drift and diffusion terms, and furthermore (ii) it also possesses the nilpotency property of (2.5). This behavior is analogous to that of linear filtering problems in which a linear signal model gives rise to the optimal filter which is also linear in both the drift and diffusion terms.

The above structural features notwithstanding, it seems unlikely that the state spaces of (2.5) and (3.7) will be <u>identical</u> nilpotent group

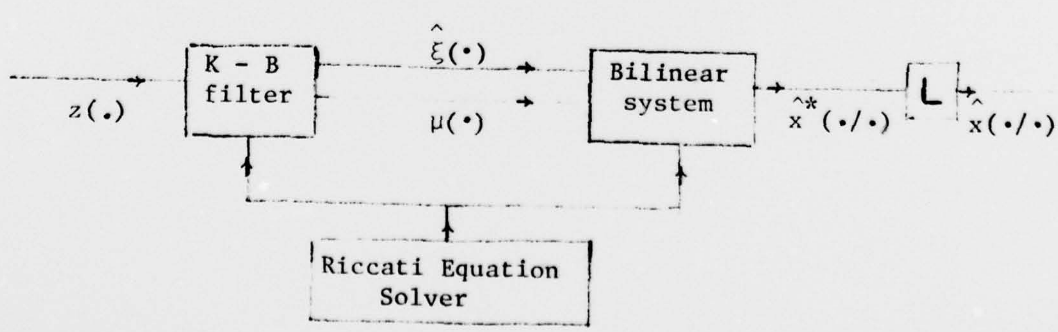
manifolds. This is obviously undesirable from a practical standpoint. One way to remedy this situation might be — rather than least-squares — to look for error criteria themselves defined on such manifolds. Such an approach for signal processes evolving on abelian Lie groups was followed in [1].

4. COMPUTATIONAL CONSIDERATIONS

Realization of the filter (3.7) in the form of a block schematic is shown below in figure 1. The practical significance of the bilinear property of the above filter is that on-line microprocessor implementation of the filter is still possible with easily available and cheap hardware consisting of integrators, summers and multipliers. This is especially important in view of the obviously huge dimensionality of this filter.

Figure 1

Block Schematic of the Optimal Nonlinear Filter



The following example illustrates the optimal filter (3.7) — for a specific choice of (2.3) — obtained by applying the alogrithm developed in the proof of theorem 5.

Example 6: $M = 3,\ N = 2,\ A = 0$

$$B_i = \begin{bmatrix} b_o^i & b_1^i & 0 \\ 0 & b_o^2 & b_2^i \\ 0 & 0 & b_o^i \end{bmatrix} , \quad i = 1,2 \quad . \tag{4.1}$$

Observe that with this choice, the system (2.3) is already in the canonical decoupled form as given by (3.11) and (3.12), and hence we may take — following the notation of the proof — $D(\cdot) \triangleq I_2$ $\zeta^*(\cdot) \triangleq \xi(\cdot)$ and $y(\cdot) \triangleq x(\cdot)$. This simplification permits a vast reduction in the filter dimensionality as follows. Since $N = N^* = 2$, $M = 3$ we have $M^* = 219$ having required three applications of Kushner equations. But in this case $D_{ij}(\cdot) \triangleq I$ so that the number of resulting augmenting states can be reduced by a factor of $2^2 = 4$, by combining them as follows. Define

$$y^k(t) = \sum_{i=j=1}^{2} y^{(k,i,j)}(t), \quad k = 1,2. \tag{4.2}$$

and

$$y^{k,k'}(t) = \sum_{i=j=1}^{2} \sum_{i'=j'=1}^{2} y^{\substack{(k,i,j) \\ (k',i',j')}}(t), \quad \ell = 1,2 \quad . \tag{4.3}$$

This yields $M^* = 21$ dimensional filter with the following coefficient matrices:

$$
{}^*(t) = \begin{bmatrix}
\sum_{\ell=1}^{2} B_\ell P_{o,\ell}(t) & B_1^1 & B_2^1 & 0 & 0 & 0 & 0 \\
\sum_{\ell=1}^{2} B_\ell P_{1\ell}(t) & \begin{array}{l}(I_3^1 \epsilon_{11}(t) \\ + \sum_{\ell=1}^{2} B_\ell^1 P_{o,\ell}(t))\end{array} & I_3^1\ \epsilon_{12}(t) & B_1^2 & 0 & B_2^2 & 0 \\
\sum_{\ell=1}^{2} {}^1B_\ell P_{2\ell}(t) & I_3^1\ \epsilon_{21}(t) & \begin{array}{l}(I_3^1\ \epsilon_{22}(t) \\ + \sum_{\ell=1}^{2} B_\ell^1 P_{o,\ell}(t))\end{array} & 0 & B_1^2 & 0 & B_2^2 \\
0 & 2\sum_{\ell=1}^{2} {}^2B_\ell P_{1\ell}(t) & 0 & \begin{array}{l}2(I_3^2\ \epsilon_{11}(t) \\ + \sum_{\ell=1}^{2} B_\ell^2\ P_{o,\ell}(t))\end{array} & I_3^2\ \epsilon_{12}(t) & 0 & 0 \\
0 & \sum_{\ell=1}^{2} {}^2B_\ell P_{2\ell}(t) & \sum_{\ell=1}^{2} B_\ell P_{1\ell}(t) & I_3^2\ \epsilon_{21}(t) & \begin{array}{l}(I_3^2\ \epsilon_{22}(t) \\ + \sum_{\ell=1}^{2} B_\ell^2 P_{o,\ell}(t))\end{array} & (I_3^2\ \epsilon_{11}(t) & I_3^2\ \epsilon_{12}(t) \\
0 & \sum_{\ell=1}^{2} {}^2B_\ell P_{2\ell}(t) & \sum_{\ell=1}^{2} {}^2B_\ell P_{1\ell}(t) & I_3^2\ \epsilon_{21}(t) & I_3^2\ \epsilon_{22}(t) & \begin{array}{l}(I_3^2\ \epsilon_{11}(t) \\ + \sum_{\ell=1}^{2} B_\ell^2 P_{o,\ell}(t)\end{array} & I_3^2\ \epsilon_{12}(t) \\
0 & 0 & 2\sum_{\ell=1}^{2} {}^2B_\ell P_{2\ell}(t) & 0 & 0 & 2I_3^2\ \epsilon_{21}(t) & \begin{array}{l}2(I_3^2\ \epsilon_{22}(t) \\ + \sum_{\ell=1}^{2} B_\ell{}^2 P_{o,\ell}(t)\end{array}
\end{bmatrix}
\quad (4.4)
$$

$$
{}^{*}_{i} = \begin{bmatrix}
B_i & & & & & & 0 \\
 & B_i^1 & & & & & \\
 & & B_i^1 & & & & \\
 & & & B_i^2 & & & \\
 & & & & B_i^2 & & \\
 & & & & & B_i^2 & \\
0 & & & & & & B_i^2
\end{bmatrix} \qquad \text{(4.5)}
$$

$i=1,2$

$$C_i^*(t) = \left[\begin{array}{c:c:c:c:c:c:c}
\sum_{\ell=1}^{2} B_\ell P_{o,\ell}(t) & I_3^1\delta(i-1) & I_3^1\delta(i-2) & 0 & 0 & 0 & 0 \\ \hdashline
0 & \sum_{\ell=1}^{2} B_\ell\,{}^1P_{o,\ell}(t) & 0 & I_3^2\delta(i-1) & 0 & I_3^2\delta(i-2) & 0 \\ \hdashline
0 & 0 & \sum_{\ell=1}^{2} B_\ell\,{}^1P_{o,\ell}(t) & 0 & I_3^2\delta(i-1) & 0 & I_3^2\delta(i-2) \\ \hdashline
0 & 0 & 0 & \sum_{\ell=1}^{2} B_\ell\,{}^2P_{o,\ell}(t) & 0 & 0 & 0 \\ \hdashline
0 & 0 & 0 & 0 & \sum_{\ell=1}^{2} B_\ell\,{}^2P_{o,\ell}(t) & 0 & 0 \\ \hdashline
0 & 0 & 0 & 0 & 0 & \sum_{\ell=1}^{2} B_\ell\,{}^2P_{o,\ell}(t) & 0 \\ \hdashline
0 & 0 & 0 & 0 & 0 & 0 & \sum_{\ell=1}^{2} B_\ell\,{}^2P_{o,\ell}(t)
\end{array}\right]$$

$i = 1,2.$

and finally

$$L(t) = \left[I_3 \;\middle|\; 0 \right] \quad : \quad 3 \times 21 \qquad (4.7)$$

where, the notation used is as follows:

$P_{ij}(t)$, $\epsilon_{ij}(t)$: i,j^{th} elements of matrices $P(t)$ and $[F(t) - P(t)H^T(t)R^{-(t)}H(t)]$ respectively

$\delta(\cdot)$: Krönecker delta function,

and for any $n \times n$ matrix U,

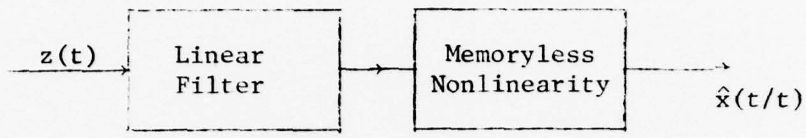
$${}^{i}U \triangleq T_i^n U \quad \text{and} \quad U^i \triangleq UT_i^n \qquad (4.8)$$

Observe the repetition of the 5^{th} and 6^{th} block rows as well as the all zero row numbers 6, 9, 11, 12, 14, 15, 17, 18, 20, 21 in the above example, so that the filter dimensionality is in effect reduced to 10. This observation can be generalized in a straightforward way, and it follows that the filter dimension in effect can be reduced to

$$\sum_{i=0}^{M-1} \binom{(N^2N^*) + i - 1}{i} \left(M-i\right) .$$

Despite the possibility of improvements of the above type, the practical implementation of such filter may at times prove formidable. We conclude this paper by pointing out three nontrivial cases in which the bilinear filter presented here collapses into an easily implementable, nonlinear, memoryless postprocessor, as shown in figure 2.

Figure 2



(i) Systems with Output Nonlinearity:

Suppose the components of the $x(\cdot)$ process to be estimated are multilinear forms in the components of a linear system driven by the $\xi(\cdot)$ process of (2.1). Systems of this type are frequently of interest in realization theory (see e.g. [18]) as they serve as good models for a wide class of nonlinear processes. Since all (conditional) moments of a multivariate gaussian distribution are completely determined by its (conditional) mean vector and (nonrandom conditional) covariance matrix, it is easy to see that $\hat{x}(t/t)$, $t \geq 0$ can be obtained as in figure (2). It can also be checked by direct differentiation that the $x(\cdot)$ process satisfies a bilinear dynamical equation with the nilpotent Lie algebra as in theorem (5).

(ii) Abelian Systems:

Suppose that the matrices $A, B_1, \ldots, B_N$ in (2.3) commute. It is easily verified that

$$x(t) = e^{\sum_{i=1}^{N} B_i \int_0^t \xi_i(\tau)d\tau} \cdot e^{At}\, x(0). \tag{4.9}$$

The desired filter structure now follows upon utilizing the gaussian characteristic function formula. (See also [1] and [19].)

(iii) Single Input Systems:

In (2.3) let $N = 1$ and let $\{A,B\}_L$ be nilpotent. If we use the canonical form of Lemma 1, it is not difficult to see that each component of $x(\cdot)$ can be written as a finite sum of terms of the form

$$y_o e^{at} \cdot e^{b\int_0^t \xi(\tau)d\tau} \int_0^t \int_0^{\sigma_1} \cdots \int_0^{\sigma_{\ell-1}} \xi(\sigma_1)\xi(\sigma_2)\cdots \xi(\sigma_\ell)d\sigma_1 \cdots d\sigma_\ell ,$$

where y_o is a random variable independent of $\xi(\cdot)$. But the ℓ fold integral in the above expression can be replaced by $\frac{1}{\ell!}\left(\int_o^t \xi(\sigma)d\sigma\right)^\ell$. Thus case (iii) is roughly a combination of cases (i) and (ii).

For the sake of completeness, we record some formulas useful in solving the above three cases. Let $Y(t) \triangleq \int_o^t \xi(\tau)d\tau$. Then,

$$E\left[e^{Y(t)}Y^\ell(t)/z^t\right] = e^{(\hat{Y}(t/t) + \frac{1}{2}\sigma^2(t))} \cdot m_\ell\left[(\hat{Y}(t/t) + \sigma^2(t), \sigma^2(t)\right] \tag{4.10}$$

where $\sigma^2(t)$: Nonrandom error covariance (computed via a Riccati equation) and $m_\ell(\eta,v)$: ℓ^{th} moment of a gaussian random variable with mean η and variance v.

Furthermore, $m_\ell(\eta, v)$, $\ell = 2,4,6,\ldots$ may be recursively computed via, (see e.g. [20], pp. 159-162)

$$\frac{\partial m_\ell(\eta,v)}{\partial v} = \frac{\ell(\ell-1)}{2} m_{\ell-2}(\eta,v) \tag{4.11}$$

with

$$m_\ell(\eta, 0) = \eta^\ell .$$

REFERENCES

1. James T. Lo and Alan S. Willsky, "Estimation for Rotational Processes with one Degree of Freedom - Part I: Introduction and Continuous Time Processes", IEEE Transactions on Automatic Control, 20, 10-21 (1975).

2. James T. Lo, "Signal Detection for Bilinear Systems", Information Sciences, 9, 249-278 (1975).

3. Steven I. Marcus and Alan S. Willsky, "Algebraic Structure and Finite Dimensional Nonlinear Estimation", SIAM Journal on Mathematical Analysis, 9, 312-327 (1978).

4. James T. Lo, "Estimation and Detection on Lie Groups", (to appear in) Nonlinear Estimation and Filtering Theory: A Status Review, E. B. Stear ed., Marcel Dekker (1979).

5. Harold J. Kushner, "Dynamical Equations for Optimal Nonlinear Filtering", Journal of Differential Equations, 3, 179-190 (1967).

6. A. H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York (1970).

7. A. V. Balakrishnan, "A General Theory of Nonlinear Estimation Problems in Control Systems", Journal of Mathematical Analysis and Applications, 8, 4-30 (1964).

8 Steel T. Huang, Nonlinear Analysis of Spherically Invariant Processes and Its Ramifications, Ph.D. Thesis, Department of Statistics, University of North Carolina at Chapel Hill (1975).

9. M. Fliess, "Un Outil Algebrique: Les Series Formelles Non Commutatives", Lecture Notes in Mathematical Systems and Economics, Springer-Verlag, Berlin (Proceedings of the Conference on Mathematical Systems Theory, Udine, Italy, 1975).

10. H. Sussman, "Semigroup Representation, Bilinear Approximation of Input Output Maps, and Generalized Inputs", Lecture Notes in Mathematical Systems and Economics, Springer Verlag, Berlin (Proceedings of the Conference on Mathematical Systems Theory, Udine, Italy, 1975).

11. Roger W. Brockett, "System Theory on Group Manifolds and Coset Spaces", SIAM Journal on Control, 10, 265-284 (1972).

12. A. A. Sagle and R.E. Walde, Introduction to Lie Groups and Lie Algebras, Academic Press, New York (1973).

13. Shirish D. Chikte and James T. Lo, "Optimal Filters for Nilpotent Associate-Algebraic Bilinear Systems", Proceedings of the Sixteenth Allerton Conference on Communication, Control and Computing (1978).

14. J. Wei and E. Norman, "On the Global Representations of the Solutions of Linear Differential Equations as a Product of Exponentials", Proceedings of the American Mathematical Society, 15, 327-334 (1964).

15. K. S. Miller, Multidimensional Gaussian Distributions, John Wiley and Sons, New York (1965).

16. R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory", Transactions of ASME, Ser. D; Journal of Basic Engineering, 83, 95-107 (1961).

17. T. Kailath and P. Frost, "An Innovations Approach to Least-Squares Estimation: Part II: Linear Smoothing in Additive White Noise", IEEE Transactions on Automatic Control, 13, 655-660 (1968).

18. Thurman R. Harper and Wilson J. Rugh, "Structural Features of Factorable Volterra Systems", IEEE Transactions on Automatic Control, 21, 822-832 (1976).

19. A. S. Willsky, "Some Estimation Problems on Lie Groups", Geometric Methods in System Theory, D. O. Mayne and R. W. Brockett eds., D. Reidel Publishing Company, Boston (1973).

20. A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill Book Company, New York (1965).

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 78-1505	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) OPTIMAL FILTERS FOR BILINEAR SYSTEMS WITH NILPOTENT LIE ALGEBRAS		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Shirish D. Chikte and James Ting-Ho Lo		8. CONTRACT OR GRANT NUMBER(s) AFOSR 74-2671
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Maryland, Baltimore County Department of Mathematics Baltimore, Maryland 21228		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE October 1978
		13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Least squares estimation, Bilinear Systems, Nilpotent Lie algebra, Linear Ito equation, Brownian motion, Additive noise, Nonlinear filter structure, Innovations.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We consider a bilinear signal process driven by a Gauss-Markov process which is observed in additive, white, gaussian noise. An exact stochastic differential equation for the least squares filter is derived when the Lit algebra associated with the signal process is nilpotent. It is shown that the filter is also bilinear and moreover that it satisfies an analogous nilpotency condition. Finally, some special cases and an example are discussed, indicating ways of reducing the filter dimensionality.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)